

얼굴 특징점의 지각적 위계구조 분석

반 세 범, 정 찬 섭

연세대학교 심리학과

Analysis of Perceptual Hierarchy for Facial Feature Point

Sebum Ban, Chansup Chung

Dept. of Psychology, Yonsei University

Abstract

표정인식 시스템을 구현하기 위해서는 어떠한 얼굴 특징점이 특정한 내적상태와 밀접한 관련이 있는가를 알아야한다. 이를 위해 MPEG-4 FDP 중 39개의 얼굴 특징점을 사용하여 쾌-불쾌 및 각성-수면의 내적상태와 얼굴 특징요소간의 상관관계를 분석하였다. 연극배우들의 다양한 표정연기 사진 150장으로부터, 5개의 필터 크기와 8개의 필터 방위로 구성된 Gabor wavelet을 사용하여 39개의 특징점을 중심으로 영상처리 하였다. 이들 특징점의 필터 반응 값과 내적상태의 상관관계를 분석한 결과, 내적상태의 쾌-불쾌 차원은 주로 입과 눈썹 주변의 특징점과 밀접한 관련이 있었고, 각성-수면 차원은 주로 눈 주변의 특징점과 밀접한 관련이 있었다. 필터의 크기는 주로 저역 공간빈도 필터가 내적상태와 관련이 있었고, 필터의 방위는 주로 비스듬한 사선 방위가 내적상태와 관련이 있었다.

Keywords: 표정인식, 얼굴인식, 내적상태, 감성 인터페이스, 감성 컴퓨터

1. 서론

표정인식 시스템을 구현하기 위해서는 어떠한 얼굴 특징점이 특정한 내적상태와 밀접한 관련이 있는가를 알아낸 후, 가장 적절한 얼굴 특징점을 선정해야 한다. 일반적으로 얼굴 및 표정인식을 위해서 눈, 눈썹, 코, 입 등에서 추출한 20개에서 40개의 특징점이 사용되고 있다. 그러나 이들 특징점들은 연구자의 기준에 따라 자의적으로 선정된 것이므로 공통적인 표준이 필요하다. 화상통신 분야에서 표준으로 제안중인 MPEG-4에서 규정된 얼굴 정의 파라미터(FDP) 및 얼굴 애니메이션 파라미터(FAP)는 얼굴표정을 결정하는 특징점의 선정에 적합한

구조를 지니고 있다. 이들 특징점 중에서 상대적 중요도를 고려하여 표정인식에 결정적인 점들을 중심으로 위계적으로 Gabor 필터를 적용하면, 작은 수의 특징점으로도 충분히 표정정보를 표상할 수 있다.

2. FDP 기반 특징점의 선정

FDP 특징점 중 우선 39개를 선정하여 표정변화가 일어났을 때의 위치이동 정보를 기록하였다. 이는 앞으로 표정인식과 표정합성을 위한 기초자료로 사용될 수 있다. 이 중 코의 가장 높은 점(그림 1의 17)은 기준점으로 사용하였으므로 거리나 각도의 변화가 없는 것으로

가정하였다.

선정된 FAP 특징점 39개 모두 표정인식에서 동일한 정도의 중요성을 지니고 있지는 않으므로, 이들 중 상대적 중요도가 높은 것과 낮은 것을 구별해 볼 필요가 있다. 특히 표정과 직접 관련이 없는 머리끝, 턱끝, 귀끝 등의 얼굴 윤곽점들을 제외해도 표정인식에는 큰 영향을 미치지 못한다. 뺨이나 턱과 관련된 특징점은 윤곽이 뚜렷하지 않으므로 위치지정이 어려우며, 주변의 다른 특징점들과 함께 움직이는 경향이 있다.

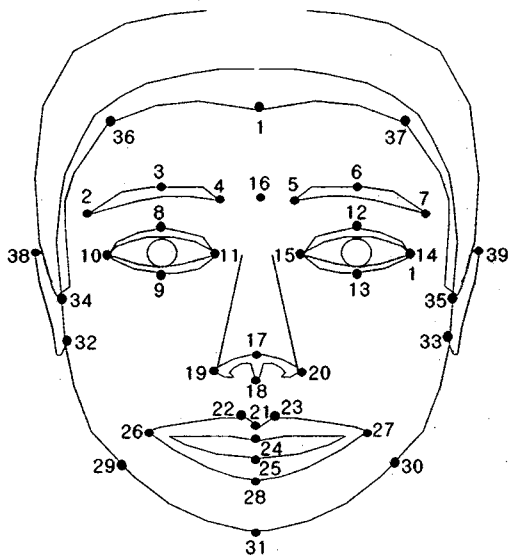


그림 1. FDP 중 선정된 39개의 특징점

이들 중 좀 더 중요한 특징점들을 골라보면 다음과 같다(그림 1). 눈썹의 특징점 중에서는 양미간 사이의 두 점(그림의 4, 5)과 중간점(그림의 16)이 특히 중요하다. 이 점들은 추미근(corrugator)과 관련되어 눈썹을 찌푸리는 것을 알 수 있게 하며, 따라서 불쾌한 내적상태와 밀접한 관련이 있다. 입술의 좌우 끝점(그림의 26, 27)은 관골근(zygomatic major)과 관련되어 미소짓는 것을 알 수 있게 하며, 유쾌한 내적상태와 밀접한 관련이 있다(Smith와 Scott, 1997). 눈의 아래위 끝점(그림의 8, 9, 12, 13)과 입술의 아래위 끝점(21, 28)은 표정의 세로

요소로서 상하운동과 관련되는데, 특히 각성상태와 밀접한 관련이 있다(한재현, 1997). 코는 얼굴의 특징요소 중 상대적으로 표정과 관련이 없지만, 코의 양끝점(19, 20)은 얼굴을 찡그리는 것을 알 수 있게 하며, 혐오나 경멸 등의 내적상태와 관련이 있다.

이와 같이 FDP 특징점에 대한 상대적 중요도 분포를 고려하면 몇 개의 특징점을 선정하는 것이 표정인식에 적절한가를 결정할 수 있다.

3. Gabor wavelet 변환과 최적화

3.1 Gabor wavelet 변환

영상에서 한 화소의 주변 영역은 동일한 화소 위치에서 서로 다른 공간 주파수와 방위의 Gabor 웨이블릿의 집합으로 표상될 수 있는데, 이러한 표상을 나타내는 집합을 제트(jet)라고 한다. 즉, 영상 $I(\vec{x})$ 에서 주어진 화소 $\vec{x} = (x, y)$ 의 위치에 다중해상 다중방위의 Gabor 웨이블릿 필터를 convolution 시키면 서로 다른 반응 값으로 구성된 제트 값 $J(\vec{x})$ 을 얻을 수 있다.

여기서 사용되는 Gabor 웨이블릿은 다음의 convolution kernel이다(Daugman, J. G., 1985; 차의영, 1998).

$$G_j(x, y) = g_j(x, y) \exp\{i\omega_j[x \cos \theta_j + y \sin \theta_j]\} \quad (\text{식 1})$$

여기서 x 와 y 는 영상의 좌표이고, ω 와 θ 는 변조 주파수와 방위이다.

가우시안 윈도우 g_j 는 다음의 식으로 만들어진다.

$$g_j(x, y) = \exp\left\{-\frac{1}{4\pi\sigma^2}[x^2 + y^2/a^2]\right\},$$

$$(x', y') = (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta) \quad (\text{식 2})$$

여기서 x' 와 y' 는 θ 만큼 회전된 방위의 영

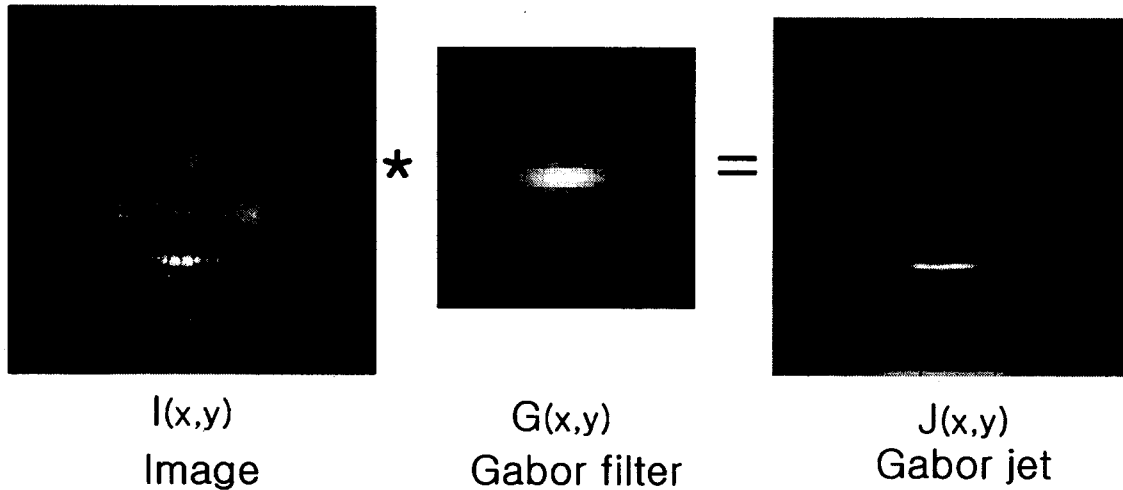


그림 2. 얼굴 영상의 Gabor wavelet 변환. 입력 영상 $I(x,y)$ 와 Gabor wavelet(여기서는 41 화소의 필터 크기와 90도 수평의 필터 방위) $G(x,y)$ 를 컨벌류션한 결과로 Gabor jet $J(x,y)$ 가 생성된다.

상좌표를 정의하며, σ 와 α 는 가우시안 원도의 크기와 이심률(eccentricity)이다.

$G_j(x,y)$ 가 영상 $I(x,y)$ 에 적용되는 Gabor함수라고 할 때, $G_j(x,y)$ 는 G_x 와 G_y 의 두 개의 분리 가능한 필터로 구성되며, 이 필터들은 90도의 위상관계를 이루고 있다. 이들 필터 쌍들의 위상차이를 최소화하기 위하여, 각각의 필터를 영상에 적용한 후 생성된 2개의 영상이 합산되어 Gabor 강도(magnitude) 값이 계산된다.

본 연구에서는 이와 같은 방법을 사용하여, 한 개의 얼굴 영상에서 39개 특징점에 대해 다중해상 다중방위로 이루어진 Gabor wavelet 제트의 강도 값으로 특징 벡터를 구성하여 얼굴 표정을 표상하였다.

영상처리에 사용된 표정자료는 6명의 연극배우로부터 수집되었는데, 쾌-불쾌와 각성-수면 차원에 골고루 분포된 83가지 내적상태에 대응되는 표정연기로 구성된 표정-내적상태 상관 DB(정찬섭, 오경자, 이일병, 변혜란, 1998)에서 총 150장을 선정하였다.

3.2. Gabor wavelet 필터구성의 최적화

Gabor wavelet을 사용하면 영상을 효율적으로 표상할 수 있다. 그러나 Gabor wavelet을 몇 개의 해상도와 방위로 구성할 것인가에 대

한 문제가 남아있다. Gabor 필터는 원래 일반적인 2차원 대상인식을 위해 만들어진 것이므로 좀 더 얼굴인식 및 표정인식에 적절한 해상도와 방위로 설정해 줄 필요가 있다.

일반적으로 필터 크기는 영상의 크기에 따라 3개에서 5개 정도로 설정하고 있다. 선행연구에 따르면 8~13 cycle/face 정도의 범위가 얼굴 정보처리에 최적이라고 한다(Nasanen, 1999).

필터 방위에 대한 선행연구를 보면 수평 방위의 필터가 최적인 것으로 나타났는데(Lyons, Budynek & Akamatsu, 1999), 이는 결국 표정에서 눈과 입의 상하 운동정보가 중요하다는 것을 의미한다. 특히 이들 정보는 각성상태와 밀접한 관계가 있다(한재현, 1997). 이에 비해 수직방위 정보는 매우 빈약하였다. Yamada와 Shibui(1998)의 연구에 따르면 표정과 관련된 얼굴구조에는 두 가지 변수가 있는데, 특징요소의 기울기(slantedness) 정보는 유쾌한 내적상태와 관련이 있으며, 곡선/열립(curvedness/openness) 정보는 각성상태와 관련이 있다고 한다.

4. 결과

4.1. FDP 특징점과 내적상태의 회귀분석

39개의 특징점 중 어떤 것이 표정인식에 결정적인가를 알아보기 위하여 내적상태의 쾌-불쾌 차원과 각성-수면 차원에 대해 회귀분석을 실시하였다. 그 결과 각 차원별로 다음 그림 3과 같이 특징점의 중요도 순서가 나타났다. 쾌-불쾌 차원의 경우 주로 입과 눈썹, 그리고 눈과 관련된 특징점의 순서로 내적상태와 밀접한 관계가 있었다. 그림에서 10% 미만의 설명력을 지닌 점들은 따로 표시하지 않았고, 괄호 안에 표시된 점들은 설명력이 떨어지지만 대칭되는 특징점들이 설명력이 높으므로 함께 표시하였다.

각성-수면 차원의 경우에는 주로 눈과 관련된 특징점들이 내적상태와 밀접한 관계가 있었다.

4.2. Gabor 필터구성과 내적상태의 회귀분석

내적상태에 대한 얼굴 특징점의 회귀분석과 함께, 어떠한 필터 크기와 필터 방위가 가장 중요한가를 알아보기 위해 회귀분석을 하였다. 39개의 얼굴 특징점에 대해 5개의 필터 크기와 8개의 필터 방위로 구성된 Gabor wavelet을 사용하였는데, 필터 크기로는 5, 10, 20, 40, 80 화소의 필터 크기를 사용하였고, 필터 방위로는 0도(수직), 22.5도, 45도, 67.5도, 90도(수평), 112.5도, 135도, 157.5도의 필터 방위를 사용하였다. 그 결과 내적상태와 밀접한 관련이 있는 필터 크기는 주로 저역 공간 주파수의 필터였고, 필터 방위로는 주로 사선 방위의 필터가 내적상태와 밀접한 관련이 있었다(표 1).

이들 결과를 바탕으로 후속연구에서 표정인식 신경망을 구성할 때, 내적상태와 밀접한 특징점과 필터 크기 및 필터 방위를 중심으로

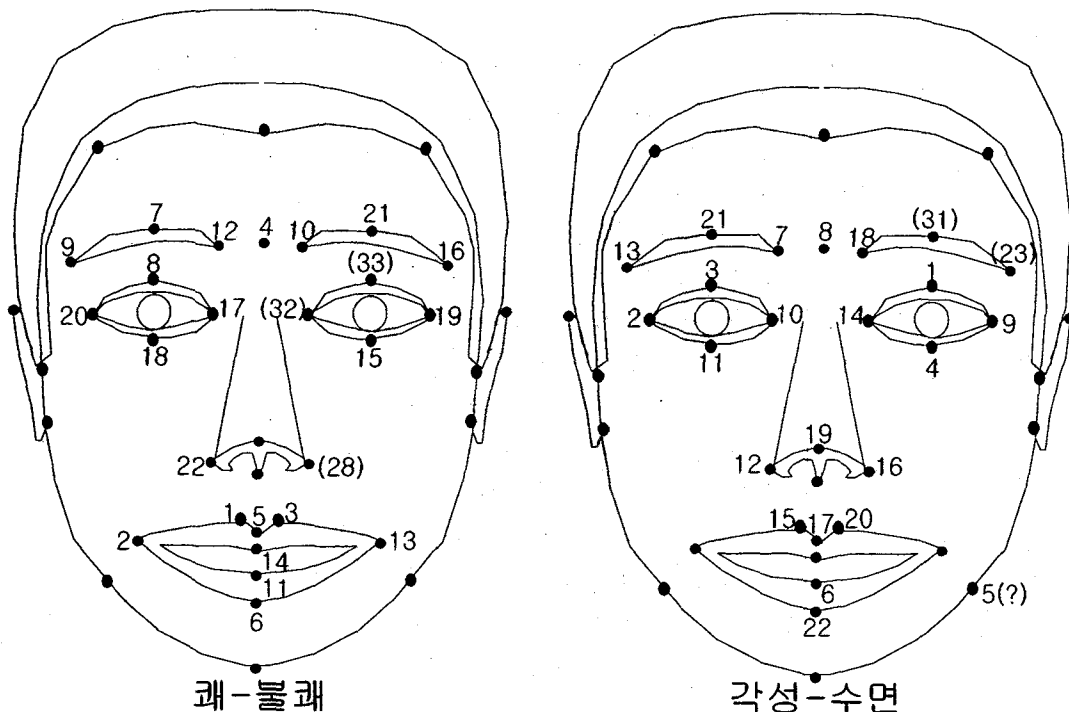


그림 3. 39개 얼굴 특징점과 내적상태의 중다 회귀분석 결과. 내적상태의 쾌-불쾌 차원은 주로 입과 눈썹주변의 특징점과 밀접한 관련이 있고, 각성-수면 차원은 주로 눈주변의 특징점과 밀접한 관련이 있다. 특징점의 숫자는 회귀분석 결과 내적상태를 예측하는 설명력이 큰 순서를 가리킨다.

표 1. 필터 크기 및 방위와 내적상태의 중다 회귀분석 결과.

패-블록		각성-수면		평균	
필터크기(화소)	설명량(R ²)	필터크기(화소)	설명량(R ²)	필터크기(화소)	설명량(R ²)
40	0.240	80	0.193	40	0.175
80	0.153	40	0.110	80	0.173
20	0.122	10	0.080	10	0.077
10	0.074	5	0.048	20	0.075
5	0.036	20	0.027	5	0.042
필터방위(각도)		필터방위(각도)		필터방위(각도)	
설명량(R ²)	설명량(R ²)	설명량(R ²)	설명량(R ²)	설명량(R ²)	설명량(R ²)
112.5	0.234	67.5	0.131	112.5	0.171
135.0	0.182	135.0	0.122	135.0	0.152
45.0	0.177	112.5	0.107	67.5	0.140
157.5	0.162	45.0	0.094	45.0	0.136
22.5	0.150	90.0	0.069	157.5	0.110
67.5	0.149	157.5	0.057	22.5	0.099
0.0(수직)	0.122	22.5	0.048	0.0	0.083
90.0(수평)	0.072	0.0	0.044	90.0	0.071

최적화된 모형을 구성하면, 보다 작은 정보로도 효율적인 신경망을 구현할 수 있다.

참고문헌

정찬섭, 오경자, 이일병, 변혜란 (1998).

표정/제스처에 의한 감성측정기술 및 DB개발. 과학기술부 감성공학기술 개발사업 최종보고서, G-17-01-09.

차의영(1998). 가보 함수와 신경회로망을 이용한 다중 목표물 탐지 및 추적. 서울대학교 대학원 컴퓨터공학과 박사학위 논문.

한재현. (1997). 얼굴표정에 의한 내적상태 추정. 연세대학교 본대학원 심리학과 석사학위논문.

Daugman, J. G. (1985) Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by 2D Visual Cortical Filters. *J. Opt. Soc. Amer. (A)* Vol. 2, pp. 1160-1169.

Lyons, M. J. , Budynek, J. , Akamatsu, S. (1999) Extracting symbolic information from facial images. 1999 Third *International Conference on Knowledge-*

Based Intelligent Information Engineering Systems, 427-430.

Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research* 39, 3824-3833.

Smith, C. and Scott, H. (1997). A componential approach to the meaning of facial expression. In Russell, J. and Fernandez-Dols, J. *The Psychology of Facial Expression*. Cambridge University Press, New York.

Yamada, H. and Shibui, S. (1998) A psychological model of recognizing facial expressions of emotion. *ATR Symposium on Face and Object Recognition '98*, 23-24.